

ISSN 1816-0301 (Print)
ISSN 2617-6963 (Online)
УДК 004.934.5:811.161.3

Паступіла ў рэдакцыю 03.01.2019
Received 03.01.2019

Прынята да публікацыі 18.01.2019
Accepted 18.01.2019

Фанетычная мінімізацыя корпуса тэкстаў на беларускай мове для навучання сістэмы сінтэзу маўлення

С. І. Лысы

*Аб'яднаны інстытут праблем інфарматыкі Нацыянальнай акадэміі навук Беларусі,
Мінск, Беларусь
E-mail: stanislau.lysy@gmail.com*

Анотацыя. Большасць сучасных сістэм сінтэзу маўлення базіруюць сваю працу на корпусным метаде. Корпусны метада, у адрозненні ад папулярнага раней кампіляцыйнага, выкарыстоўвае базу дадзеных натуральнага маўлення, якая складаецца не з асобных спецыяльна выбраных элементаў кампіляцыі, а ўяўляе сабой корпус фанэграм натуральнага маўлення. Для дасягнення высокай якасці сінтэзаванага маўлення пры такім падыходзе патрабуюцца вялікія аб'ёмы тэкставай і адпаведнай гукавой інфармацыі, што з'яўляецца істотнай праблемай для так званых нерэсурсных моў, да якіх адносіцца і беларуская. У такім выпадку, як правіла, прымяняецца фанетычная мінімізацыя – адмысловы адбор тэкстаў, у выніку якога аб'ём тэкставага корпуса максімальна змяншаецца, але пры гэтым захоўваецца фанетычная паўната.

У артыкуле разглядаюцца звесткі пра сутнасць і спосаб працы корпуснага метада генерацыі гукавога сігнала ў сістэмах сінтэзу маўлення, прыводзіцца падрабязны агляд падыходаў да фарміравання тэкставых і маўленчых карпусоў, неабходных для генерацыі маўлення корпусным метадам. Другая палова працы прысвечана апісанню распрацаванага алгарытму фанетычнай мінімізацыі корпуса тэкстаў на беларускай мове, а таксама тэхнічных і лінгвістычных рэсурсаў, выкарыстаных для яго рэалізацыі. Прыводзіцца апісанні распрацаванага праграмага прататыпа і шэрагу праведзеных аўтарам эксперыментаў па фанетычнай мінімізацыі.

Ключавыя словы: фанетычная мінімізацыя, беларуская мова, сінтэз маўлення, корпусны метада, корпус тэкстаў

Падзякі. Даследаванне выканана пры падтрымцы гранта НАН Беларусі № 2018-25-032 «Алгарытмы інтэрнэт-сінтэзатара беларускага маўлення і аўтаматызаванага стварэння лінгвістычных рэсурсаў», а таксама ў межах праекта БРФФД № Ф17МС-039 «Высакаякасны інтэрнэт-увод і інтэрнэт-вывад маўлення, захаванне і сістэматызацыя вялікіх аб'ёмаў (Big Data) маўлення».

Для цытавання. Лысы, С. І. Фанетычная мінімізацыя корпуса тэкстаў на беларускай мове для навучання сістэмы сінтэзу маўлення / С. І. Лысы // Информатика. – 2019. – Т. 16, № 1. – С. 75–85.

Phonetic minimization of the text corpus in Belarusian for the speech synthesis system training

Stanislau I. Lysy

*The United Institute of Informatics Problems of the National Academy of Sciences of Belarus,
Minsk, Belarus
E-mail: stanislau.lysy@gmail.com*

Abstract. The most modern speech synthesis systems are based on the corpus-based method. The corpus-based method, unlike previously popular compilation method, uses natural speech database that does not consist of separate specially selected elements of compilation, but represents the corpus of phonograms of natural speech.

Large amounts of text and corresponding audio information, which represents a significant challenge for so-called under-resourced languages, which include Belarusian, are required to achieve high-quality synthesized speech in this approach. In this case, a common approach is to use phonetic minimization, special selection of texts, when the amount of text corpus is maximally reduced, but at the same time phonetic fullness is preserved. The article discusses the information about the nature and the functioning the corpus-based method of sound signal generation in speech synthesis systems, provides a detailed overview of the approaches to the formation of text and speech corpuses, required for speech generation by the corpus-based method. The second half of the work is devoted to the description of the elaborated algorithm of the text corpus phonetic minimization in Belarusian language, as well as technical and linguistic resources used to implement it. A description of the developed software prototype as well as a description of the series of experiments on phonetic minimization are given to demonstrate the efficiency of the algorithm.

Keywords: phonetic minimization, the Belarusian language, speech synthesis, corpus-based method, text corpus

Acknowledgements. This research was supported by the NASB grant № 2018-25-032 "Algorithms of Belarusian speech web-synthesizer and automated creation of linguistic resources", as well as the BRFFR project № Ф17МС-039 "Algorithms of high quality speech Internet-input and Internet-output, storage and systematization of speech BigData".

For citation. Lysy S. I. Phonetic minimization of the text corpus in Belarusian for the speech synthesis system training. *Informatics*, 2019, vol. 16, no. 2, pp. 75–85 (in Belarusian).

Уводзіны. Агляд апошніх тэндэнцый у сферы сінтэзу маўлення паказвае, што большасць сучасных сістэм базіруюць сваю працу на падыходзе, заснаваным на выкарыстанні моўных мадэляў. Гэты падыход дазваляе замяніць класічную схему сістэм сінтэзу маўлення, якая складаецца з блокаў тэкставай, фанетычнай, акустычнай і інтанацыйнай апрацоўкі, натрэніраванай нейроннай сеткай. Аднак пры такім падыходзе для дасягнення высокай якасці сінтэзаванага маўлення патрабуюцца вялікія аб'ёмы тэкставай і адпаведнай гукавой інфармацыі, што ўяўляе сабой істотную праблему для так званых нерэсурсных моў (англ. under-resourced languages), да якіх адносіцца і беларуская [1]. У сувязі з гэтым можа быць прапанавана гібрыдная структура сінтэзатара маўлення, якая аптымальна аб'ядноўвае блокі апрацоўкі, уласцівыя канкатэна-тыўнаму падыходу, з найбольш сучаснымі дасягненнямі, звязанымі са стварэннем моўных мадэляў і навучаннем нейронных сетак. Такая гібрыдная структура, да прыкладу, можа сумясціць добра прапрацаваныя блокі тэкставай [2] і фанетычнай [3] апрацоўкі, якія выкарыстоўваюць спісы правіл, з блокам акустычнай апрацоўкі, праца якога будзе заснавана на моўнай мадэлі. Гэтая мадыфікацыя ўжо сама па сабе дазволіць знізіць колькасць неабходных моўных рэсурсаў, аднак не ў дастатковай ступені.

Для далейшага зніжэння аб'ёму патрабуемых лінгваакустычных рэсурсаў неабходна здзейсніць фанетычную мінімізацыю – адмысловы адбор тэкстаў, у выніку якога аб'ём тэкставага корпуса максімальна зменшыцца, але пры гэтым будзе захавана фанетычная паўната. Гэта значыць, што выніковае мноства тэкстаў павінна захаваць усё фанетычнае багацце мэтавай мовы.

З наступленнем эпохі асабліва імклівага развіцця інфармацыйных тэхналогій працэс фанетычнай мінімізацыі (як правіла, да той ці іншай ступені) аўтаматызаваўся. У міжнароднай практыцы задача аўтаматызаванай фанетычнай мінімізацыі неаднаразова вырашалася, аднак часцей за ўсё для найбольш запатрабаваных моў, бо з'яўляецца мовазалежнай. Агляд тэматычнай літаратуры паказаў, што на сённяшні дзень не сустракаецца апісанняў якаснай аўтаматызацыі працэсу фанетычнай мінімізацыі для беларускай мовы. З гэтай нагоды было вырашана распрацаваць алгарытм фанетычнай мінімізацыі корпуса тэкстаў на беларускай мове. Дадзены алгарытм і створаны на яго аснове прататып сістэмы фанетычнай мінімізацыі корпуса тэкстаў на беларускай мове пакліканы паспрыяць у здзяйсненні наступных мэт:

– распрацоўкі новай версіі вэб-сінтэзатара маўлення лабараторыі распазнавання і сінтэзу маўлення Аб'яднанага інстытута праблем інфарматыкі Нацыянальнай акадэміі навук Беларусі з выкарыстаннем найноўшых дасягненняў у дадзенай сферы, а менавіта з выкарыстаннем моўных мадэляў (URL: <https://corpus.by/TextToSpeechSynthesizer>);

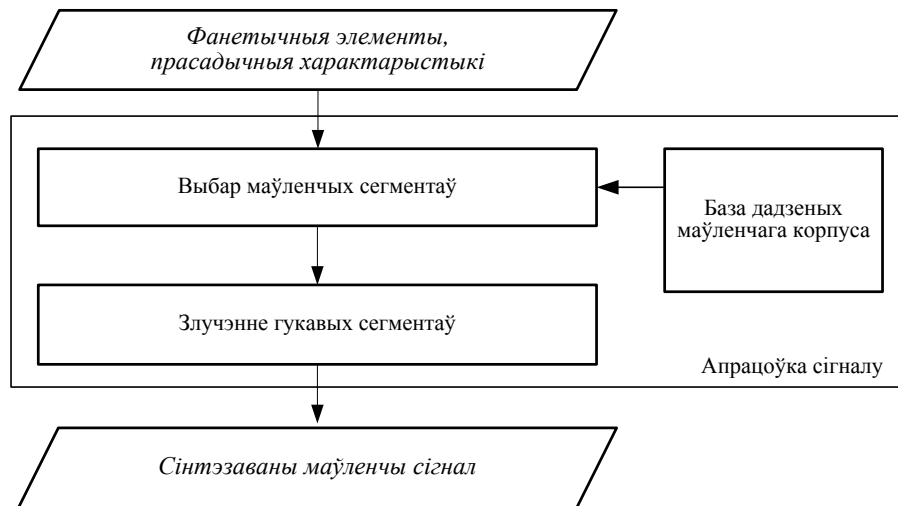
– успамогі іншым распрацоўшчыкам пры стварэнні новых сістэм сінтэзу маўлення па тэксце для беларускай мовы;

– аўтаматызацыі адбору мінімізаванага фанетычна поўнага мноства тэкстаў на беларускай мове для выкарыстання ў шэрагу разнастайных навуковых сфер, напрыклад у лінгвістычных даследаваннях ці пры стварэнні адмысловых дапаможнікаў па вывучэнні беларускай фанетыкі.

Корпусны метада сінтэзу маўлення. На працягу апошніх 50 гадоў у сістэмах сінтэзу маўлення ўжывалася мноства тэхналогій генерацыі маўлення. Адным з распаўсюджаных сучасных падыходаў з’яўляецца корпусны метада.

Корпусны метада таксама, як і кампіляцыйны, выкарыстоўвае базу дадзеных (БД) натуральнага маўлення. Аднак згаданая БД складаецца не з асобных спецыяльна выбраных элементаў кампіляцыі, а ўяўляе сабой корпус фанаграм натуральнага маўлення, размечанага на элементы фанемнай размернасці з маркерамі іх прасадыхных характарыстык. Пры гэтым важным адрозненнем корпуснага падыходу з’яўляецца таксама магчымасць выкарыстання некалькіх сегментаў з аднолькавымі фанетычнымі, але рознымі прасадыхнымі характарыстыкамі [4], дзякуючы чаму натуральны маўленчы сігнал у працэсе сінтэзу ўва многіх выпадках не падвяргаецца дадатковай прасадыхнай мадыфікацыі. Усё ж такая магчымасць не выключаецца і пры выкарыстанні кампіляцыйнага метаду ў выпадку, калі ў БД сінтэзатара будуць дубліравацца аднолькавыя фанетычныя сегменты, але з рознымі прасадыхнымі характарыстыкамі [5].

Схема корпуснага метаду сінтэзу маўлення па тэксце паказана на мал. 1 [6]. Паслядоўнасць фанетычных элементаў і адпаведныя мэтавыя прасадыхныя параметры падаюцца ў блок выбару маўленчых сегментаў, які звяртаецца да БД, што змяшчае маўленчы корпус, папярэдне размечаны на акустыка-фанетычных элементах з маркерамі прасадыхных параметраў. З БД выбіраюцца сегменты, найбольш блізкія да неабходных і па фанетычных, і па прасадыхных параметрах, якія затым часцяком без усялякай дадатковай мадыфікацыі аб’ядноўваюцца ў безупынным маўленчы сігнал [5].



Мал. 1. Схема корпуснага метаду генерацыі маўлення

Разглядаючы корпусны метада сінтэзу маўлення як сродак перадачы індывідуальных характарыстык голасу і дыкцыі асобы, неабходна адзначыць яго перавагі ў параўнанні з іншымі метадамі: выкарыстанне сегментаў натуральнага маўлення і адсутнасць мадыфікацыі маўленчага сігналу, якія гарантуюць перадачу ўсіх індывідуальных акустычных характарыстык пры сінтэзе маўлення. З іншага боку, ужыванне толькі сегментаў натуральнага маўлення накладвае пэўныя патрабаванні на склад дзеючай маўленчай базы дадзеных. Для перадачы індывідуальных фанетычных і прасадыхных характарыстык маўлення БД мусіць змяшчаць максімальную колькасць усіх магчымых камбінацый фанетычных элементаў, генеруемых фанетычнай мадэллю,

з разнастайнымі інтанацыйнымі характарыстыкамі ад прасадычнай мадэлі. Стварэнне такой БД звязана з вялікімі складанасцямі, а яе аб'ём можа стаць празмерна вялікім, што непрымальна для шэрагу практычных прыкладанняў [5].

Падыходы да фарміравання тэкставых і маўленчых карпусоў. Абгрунтаванае фарміраванне тэкставага і маўленчага карпусоў разам з лінгвістычна абгрунтаванымі класіфікацыяй і выбарам базавых сегментаў для сінтэзу маўлення шмат у чым вызначаюць ступень разборлівасці і натуральнасці сінтэзуемага маўлення. Фарміраванне карпусоў паводле зададзеных базавых сегментаў акрамя асноўнай мэты, якой з'яўляецца пакрыццё ўсіх сегментаў, павінна забяспечыць [7]:

- фанетычную паўнату выніковага корпуса (у фанетычнай транскрыпцыі тэксту мусяць сустракацца ўсе асноўныя фанемы маўлення і іх варыянты);

- па магчымасці, мінімізаваны аб'ём корпуса;

- фанетычную збалансаванасць створанага корпуса (размеркаванне частот сустракаемасці фанем і іншых фанетычных адзінак у сфарміраваным корпусе мусіць быць бліжкім да тэарэтычнага, атрыманага на дастаткова прадстаўнічых і вялікіх па аб'ёме выбарках).

Апошняе патрабаванне да фарміравання карпусоў прад'яўляецца ў выпадку, калі пры выбары базавых сегментаў не ўлічваліся фанетычныя кантэксты.

Адзін з падыходаў да фарміравання тэкставых карпусоў – экспертны падыход – заключаецца ў падборы слоў ці фраз, складаючых корпус, дасведчанымі экспертамі-фанетыстамі «ўручную» [8] або паўаўтаматычна (з выкарыстаннем праграм-транскрайбераў). Перавагі такога падыходу заключаюцца ў гарантваным пакрыцці ўсяго неабходнага мноства і ў падборы найбольш падыходзячых для вымаўлення слоў і фраз. Экспертны падыход пры пэўных умовах гарантуе атрыманне мінімальнага ці блізкага да мінімальнага аб'ёму корпуса. Да недахопаў дадзенага падыходу можна аднесці вялікі аб'ём ручной працы і патрабаванні высокай кваліфікацыі да эксперта-фанетыста, які яе выконвае.

Сярод іншых спосабаў стварэння тэкставай БД варта адзначыць падыход, які, у адрозненні ад традыцыйных падыходаў, у якасці тэкставага корпуса выкарыстоўвае бессэнсоўныя словы, або псеўдасловы. Перавагамі такога метаду з'яўляецца тое, што не патрабуецца здзяйсняць адмысловую выбарку на аснове тэкставага корпуса, а таксама ёсць гарантыя мінімальнасці аб'ёму створанай маўленчай базы, паколькі псеўдасловы могуць фарміравацца акурат так, каб утрымліваць толькі базавыя сегменты ў мінімальнай іх колькасці. Да недахопаў, аднак, можна аднесці складанасць начыткі дыктарамі псеўдаслоў, што істотна зніжае якасць запісанага маўлення, і прасадычную непрадказальнасць, што таксама вядзе да страты якасці маўленчага корпуса [5].

Пры аўтаматычным статыстычным падыходзе фарміраванне тэкставага корпуса адбываецца на аснове набору тэкстаў розных жанраў [9, 10], уключаючы газетныя артыкулы, навуковыя тэксты, мастацкія творы. Пры наяўнасці толькі фанаграм маўлення ў якасці зыходнага набору тэкстаў выкарыстоўваюцца стэнаграмы запісаў [11]. Мяркуюцца, што стэнаграмы тэкстаў вялікага аб'ёму з высокай ступенню верагоднасці ўключаюць усе магчымыя ў мове сегменты маўлення. Далей з набору гэтых тэкстаў здзяйсняецца выбар слоў ці фраз, а ўжо з іх фарміруецца новы тэкставы корпус, які мусіць быць мінімальнага памеру і пакрываць усё мноства базавых сегментаў для сінтэзу маўлення. Рашэнне гэтай задачы фармулюецца як рашэнне задачы мінімізацыі пакрыцця мноства.

Задача мінімізацыі пакрыцця мноства з'яўляецца NP-складанай [12], і для яе рашэння неабходны розныя аптымізацыйныя алгарытмы, у прыватнасці так званы паглынаючы (сквапны) алгарытм [13] і разнастайныя яго мадыфікацыі, а таксама генетычны алгарытм [14]. Сутнасць сквапнага алгарытму заключаецца ў вызначэнні найбольшай колькасці базавых сегментаў у кожным элеменце тэксту (слове, сказе, інш.) і дабаўленні на кожным кроку алгарытму ў фарміруемы новы тэкставы корпус элементаў тэксту, якія змяшчаюць найбольшую колькасць патрабуемых сегментаў. Працэс дабаўлення элементаў працягваецца да таго часу, пакуль фарміруемы тэкставы корпус не будзе ўключаць усе неабходныя сегменты. У некаторых сістэмах у сфарміраваны корпус уручную дадаюцца элементы, якія ўтрымліваюць рэдка [9], ці,

наадварот, найбільш сустракаемыя сегменты [10]. Такі спосаб фактычна з'яўляецца камбінацыяй экспертнага і статыстычнага падыходаў.

Пры выкарыстанні генетычнага алгарытму выбіраецца пачатковы набор сказаў (індывідаў), які пакрывае ўсе базавыя элементы (зыходную папуляцыю). У якасці такога набору могуць быць выкарыстаны ўсе сказы, якія складаюць зыходны корпус. Ацэнка сказа адбываецца на аснове колькасці базавых сегментаў, якія ў ім прадстаўлены. Затым ітэратыўна здзяйсняюцца аперацыі мутацыі індывідаў, іх скрыжавання, селекцыі і адбору найбольш элітных, прадстаўнічых. Пасля кожнай ітэрацыі фарміруецца новае пакаленне з лепшымі характарыстыкамі, г. зн. фарміруецца падмноства сказаў, магутнасць якога менш за магутнасць падмноства на мінулае ітэрацыі, але сфарміраванае падмноства пакрывае набор базавых сегментаў.

Ва ўсіх падыходах, якія прымяняюцца для рашэння задачы далейшай мінімізацыі набору базавых сегментаў, пастуліруецца наступнае:

- зыходны корпус ужо пакрывае ўсе неабходныя для сінтэзу базавыя сегменты;
- у корпусе прысутнічаюць элементы, празмерныя з пункту гледжання фанетыкі і прасодыі;
- у корпусе прысутнічаюць элементы, якія з-за сваіх нестандартных акустычных і пра-содычных характарыстык ніколі не будуць выбраны для сінтэзу.

На аснове двух апошніх пастулатаў і адбываецца змяншэнне маўленчай базы [5].

Параметры, выкарыстоўваемыя пры мінімізацыі тэкставых карпусоў. Перш за ўсё варта адзначыць, што адмысловы падбор сказаў для стварэння маўленчага корпуса не заўсёды прадстаўляецца магчымым. Падобная сітуацыя, напрыклад, можа ўзнікнуць, калі маўленчы корпус не адмыслова фарміруецца і запісваецца, а бяраецца з гатовых вытокаў, такіх як радыётрансляцыі, аўдыякнігі і інш. [15].

Пры падрыхтоўцы ж корпуса з нуля перад яго распрацоўшчыкамі паўстае не толькі пытанне распрацоўкі метадаў адбору збалансаванай мінімізаванай базы тэкстаў, але і пытанне выбару базавых фанетычных адзінак, на аснове якіх вядзецца адбор, а таксама пытанне вызначэння мяжы пошуку кожнай унікальнай фанетычнай адзінкі і інш.

У якасці базавай фанетычнай адзінкі часцей за ўсё прымяняюцца так званыя *n*-фоны, якія ўяўляюць сабой паслядоўнасць з *n* гукаў [16]. Прыватнымі выпадкамі прымяняемых *n*-фонаў з'яўляюцца алафоны, дыфоны і трыфоны – паслядоўнасці з аднаго, двух ці трох гукаў адпаведна. Сярод іншых магчымых варыянтаў базавай фанетычнай адзінкі можна адзначыць склады і словы [17, 18].

Што тычыцца аб'ёму, да якога варта мінімізаваць тэкставы корпус, то ён істотна залежыць ад таго, якой якасці сінтэзатар павінен атрымацца ў выніку, а таксама ад рэсурсаў, якія маюць у сваім распараджэнні распрацоўшчыкі. Так, стваральнікі галасоў для каталонскага сінтэзу маўлення ў артыкуле [19] адзначаюць, што для дасягнення высокай якасці створаных галасоў яны запісалі маўленне двух дыктараў аб'ёмам 10 г (каля 90 тыс. слоў) кожны. Для дасягнення ж сярэдняй якасці сінтэзу выконвалася скарачэнне базы да 1 г маўлення. Аўтары іншага даследавання пры стварэнні іспанскіх галасоў таксама арыентаваліся на дзесяцігадзінны аб'ём маўленчага корпуса [20].

Варта адзначыць, што ў якасці вытоку вялікага аб'ёму тэкстаў, на аснове якіх фарміруецца корпус, нярэдка выкарыстоўваецца база тэкстаў Вікіпедыі той ці іншай мовы, так як яна з'яўляецца свабоднай і не звязана аніякімі аўтарскімі правамі. Такая база, напрыклад, лягла ў аснову голасу на мове эсперанта ў сінтэзатары RHVoice. Стваральнікі гэтага голасу адзначаюць, што аўтаматычным чынам адабралі сказы памерам ад 5 да 20 слоў без лічбаў і сімвалаў, патрабуючых расшыфроўкі. Вынікам стаў корпус у 50 тыс. сказаў, на аснове якога ствараўся аптымальны мінімізаваны набор аб'ёмам 1100 сказаў (паводле стэнаграмы інтэрв'ю з распрацоўшчыцай RHVoice В. Якаўлевай; URL: https://groups.google.com/forum/#!topic/bel-voice/Sr6pPFLw_84).

У дысертацыі А. І. Саломеннік «Ацэнка якасці селектыўнага сінтэзу маўлення: метады і вынікі» (URL: https://www.philol.msu.ru/~ref/dcx/2016_SolomennikAI_avtoreferat_10.02.21_24.pdf) прыводзіцца наступнае абагульненне: «Зазвычай аб'ём маўленчых карпусоў для селектыўнага сінтэзу маўлення вагаецца ад 2 да 15 гадзін маўлення аднаго дыктара. Пры стварэнні маўленчага корпуса могуць выкарыстоўвацца спецыяльна падрыхтаваныя фанетычна

прадстаўнічыя тэксты, спісы частотных слоў, лікі, абрэвіятуры, тэксты з улікам вобласці выкарыстання сінтэзаванага маўлення».

Алгарытм фарміравання мінімізаванага корпуса тэкстаў. Прапанаваны ніжэй алгарытм дае магчымасць сфарміраваць мінімізаваны корпус тэкстаў, якія пакрываюць усе прысутныя ў зыходным корпусе гукавыя адзінкі.

Паколькі алгарытм быў першапачаткова арыентаваны на інтэграцыю з сінтэзатарам маўлення (URL: <https://corpus.by/TextToSpeechSynthesizer>), то ў якасці фанетычнага прадстаўлення тэксту браўся алафонны тэкст, згенераваны дадзеным сінтэзатарам. Такі тэкст уяўляе сабой паслядоўнасць абазначэнняў алафонаў, паўз, словападзелаў і складападзелаў. Ніжэй прыведзены фрагмент арфаграфічнага тэксту і адпаведны яму алафонны тэкст [3].

Груша цвіла апошні год. Усе галіны яе, усе вялікія расохі, да апошняга пруючка, былі ўсыпаны бурным бела-ружовым цветам.

```
GH004,R022,U022,>,SH002,A323,/,>,C'002,V'002,I241,>,L002,A012,/,>,A221,>,P001,O012,>,
SH002,N'004,I242,/,>,GH001,O032,T000,/,>,#P4,>,U203,>,S'001,E042,/,>,GH004,A233,>,L'002,
I042,>,N004,Y323,/,>,J'012,A243,>,J'011,E040,/,>,#C3,>,U203,>,S'001,E043,/,>,V'012,A243,>,
L'002,I043,>,K'002,I343,>,J'012,A342,/,>,R002,A222,>,S001,O023,>,H'002,I340,/,>,#C3,>,D004,
A322,>,A221,>,P001,O012,>,SH002,N'004,A342,>,GH004,A231,/,>,P002,R012,U023,>,C'002,I342,
>,K004,A330,/,>,#C3,>,B002,Y013,>,L'004,I241,/,>,W013,S001,Y021,>,P002,A312,>,N004,Y221,/
>,B002,U012,R001,>,N004,Y221,M001,/,>,B'002,E141,>,L004,A312,>,R002,U222,>,ZH002,O021,
>,V012,Y211,M003,/,>,C'002,V'001,E042,>,T002,A321,M000,/,>,#P4
```

Коды алафонаў у гэтым запісе складаюцца з літарнай назвы фанемы (напрыклад, GH), знака мяккасці «'» (пры яе наяўнасці) і лічбавага кода, які ўказвае на тыя ці іншыя асаблівасці фанемы. У шэрагу эксперыментаў, апісаных ніжэй, будучы выкарыстоўвацца не толькі поўныя (напрыклад, ZH002), але і скарачаныя (напрыклад, ZH0) запісы алафонаў, у якіх адкідаюцца дзве апошнія лічбы, што ўказваюць на кантэкст алафона ў слове. Таксама ў дадзеным запісе можна назіраць знакі словападзелу «/» і складападзелу «>».

Уваходныя дадзеныя алгарытму:

- файл з корпусам сказаў у алафонным запісе, F_s ;
- базавая фанетычная адзінка, *unit*;
- мяжа пошуку, *lim* (вызначае мінімальную неабходную колькасць кожнай унікальнай фанетычнай адзінкі).

Пачатак алгарытму.

Крок 1. Фарміраванне спісу сказаў. Выконваецца загрузка файла са спісам сказаў у алафонным запісе F_s . Фарміруецца спіс $L_s = \langle s_1, \dots, s_N \rangle$, дзе s_n – n -ы сказ, $n = 1, \dots, N$.

Крок 2. Падбор сказаў з неабходнымі фанетычнымі адзінкамі. Ствараюцца пустое мноства L_s' для занясення ў яго абраных сказаў, а таксама лічылнік фанетычных адзінак Cnt , дзе збіраюцца адпаведнасці фанетычных адзінак і іх колькасцяў у абраных сказах мноства L_s' . Кожны сказ s_n з мноства сказаў L_s праходзіць крокі 2.1–2.3.

Крок 2.1. Фарміраванне спісу фанетычных адзінак сказа. Вылучаюцца ўсе фанетычныя элементы сказа s_n паводле абранай базавай фанетычнай адзінкі *unit*. Калі базавай фанетычнай адзінкай з'яўляецца алафон або склад, то падзел выконваецца па адпаведных раздзяляльных сімвалах, прысутных у алафонным запісе сказаў. У выпадку, калі базавай адзінкай з'яўляюцца камбінацыі вышэй названых адзінак, такія як дыфон, трыфон і інш., спярша адбываецца падзел на элементарныя адзінкі, з якіх потым паслядоўна складаюцца камбінацыі. Вынікам будзе мноства фанетычных адзінак $U_n = \langle u_1, \dots, u_M \rangle$, дзе u_m – m -я фанетычная адзінка сказа s_n , $m = 1, \dots, M$.

Крок 2.2. Абранне сказа. Здзяйсняецца перабір фанетычных адзінак u_m з мноства U_n . Калі чарговая фанетычная адзінка сустракаецца ў лічылніку Cnt менш за *lim* разоў, то перабір прыпыняецца і адбываецца пераход да кроку 2.3. Іначай, калі перабраны ўсе фанетычныя адзінкі сказа, але ніводная не адпавядае вышэй указанай умове, пераходзім да наступнага сказа s_{n+1} і кроку 2.1. Калі ж апрацаваны сказ з'яўляецца апошнім сказам мноства ($n = N$), то пераходзім да кроку 3.

Крок 2.3. Занясенне сказа ў мноства абраных. Выконваецца перабор фанетычных адзінак u_m з мноства U_n . Калі u_m адсутнічае ў лічылніку Cnt , то ствараецца адпаведны запіс са значэннем адзінкі. Калі ж фанетычная адзінка ўжо ёсць у лічылніку Cnt , то адпаведнае ёй значэнне павялічваецца на адзінку. Калі перабраны ўсе фанетычныя адзінкі сказа, адбываецца пераход да наступнага сказа s_{n+1} і кроку 2.1. Калі ж апрацаваны сказ з'яўляецца апошнім сказам спісу L_s ($n = N$), то здзяйсняецца пераход да кроку 3.

Крок 3. Мінімізацыя мноства абраных сказаў. Ствараецца пустое мноства L_s'' для занясення ў яго толькі тых абраных сказаў з мноства L_s' , якія ўтрымліваюць хаця б адзін фанетычны элемент, выдаленне якога зніжае агульную колькасць такіх элементаў ніжэй мяжы пошуку lim . Ствараецца копія лічылніка фанетычных адзінак $Cnt - Cnt'$, дзе збіраюцца адпаведнасці фанетычных адзінак і іх колькасцяў у мінімізаваным мностве L_s'' . Кожны сказ s'_k з мноства сказаў L_s' праходзіць наступныя чатыры крокі.

Крок 3.1. Фарміраванне спісу фанетычных адзінак сказа. Вылучаюцца ўсе фанетычныя элементы сказа s'_k паводле абранай базавай фанетычнай адзінкі $unit$ па аналогіі з крокам 2.1. Мноства фанетычных адзінак трансфармуецца ў мноства адпаведнасцяў фанетычных адзінак і іх колькасцяў у мностве. Такім чынам вынікам кроку з'яўляецца мноства $U'_k = \langle \langle u'_1, cnt_1 \rangle, \dots, \langle u'_j, cnt_j \rangle \rangle$, дзе $u'_j - j$ -я фанетычная адзінка, cnt_j – колькасць сустраканняў фанетычнай адзінкі u'_j у сказе s'_k , $j = 1, \dots, J$.

Крок 3.2. Абранне сказа. Здзяйсняецца перабор фанетычных адзінак u'_j з мноства U'_k . Калі розніца колькасці сустраканняў фанетычнай адзінкі u'_j у мностве абраных сказаў L_s' і мностве фанетычных адзінак бягучага сказа U'_k менш за мяжу lim , то перабор прыпыняецца – адбываецца пераход да кроку 3.3. Іначай, калі перабраны ўсе фанетычныя адзінкі сказа, але ніводная не адпавядае вышэй указанай умове, – адбываецца пераход да кроку 3.4.

Крок 3.3. Занясенне сказа ў мінімізаванае мноства. Сказ s'_k заносіцца у мінімізаванае мноства L_s'' . Адбываецца пераход да наступнага сказа s'_{k+1} і кроку 3.1. Калі ж апрацаваны сказ з'яўляецца апошнім сказам мноства ($k = K$), то пераходзім да кроку 4.

Крок 3.4. Выдаленне сказа. Усе фанетычныя элементы сказа s'_k прадстаўлены ў астатніх сказах мноства L_s' у дастатковай колькасці, таму сказ не ўносіцца ў мінімізаваны спіс L_s'' . Здзяйсняецца перабор фанетычных адзінак u'_j з мноства U'_k . Частоты кожнай фанетычнай адзінкі cnt_k адымаюцца ад частаты адпаведнай фанетычнай адзінкі ў лічылніку Cnt' . Калі перабраны ўсе фанетычныя адзінкі сказа, адбываецца пераход да наступнага сказа s'_{k+1} і кроку 3.1. Калі ж апрацаваны сказ з'яўляецца апошнім сказам мноства ($k = K$), то пераходзім да кроку 4.

Крок 4. Фарміраванне выніку. Мноства сказаў L_s'' сартыруецца паводле беларускага алфавіту, а лічылнік Cnt' – паводле частаты. Вынік прыводзіцца ў наглядны тэкставы фармат, выводзіцца на экран і захоўваецца ў файле на серверы.

Канец алгарытму.

Пасля выканання апісанага вышэй алгарытму на аснове корпуса тэкстаў на беларускай мове фарміруецца мінімізаванае мноства сказаў, якія пакрываюць усе фанетычныя адзінкі, наяўныя ў зыходным корпусе. Варта адзначыць, што апісаны алгарытм арыентаваны не на дасягненне абсалютнай мінімізацыі, а на значнае зніжэнне аб'ёму корпуса. У кантэксце працы над сістэмай сінтэзу маўлення важна, каб алгарытм, здзяйсняючы поўную мінімізацыю, фарміраваў выніковы набор сказаў як мага больш зручным для дыктарскай начыткі і як мага больш рэпрэзентатыўным для стварэння моўнай мадэлі не толькі з пункту гледжання фанетыкі, але таксама лексікі і прасодыкі. У сувязі з гэтым для павышэння якасці выніку працы алгарытму рэкамендуецца праводзіць папярэднюю сартыроўку сказаў зыходнага корпуса, каб першымі на апрацоўку алгарытму падаваліся сказы з найбольш частотнай лексікай і найбольш рэдкімі інтанацыйнымі канструкцыямі.

Для апрабацыі, тэставання і нагляднасці працаздольнасці апісанага вышэй алгарытму быў распрацаваны прататып сістэмы фанетычнай мінімізацыі корпуса тэкстаў на беларускай мове ў форме вэб-сэрвіса «Фанетычны мінімізатар» (URL: <https://corpus.by/PhoneticMinimizer/>; мал. 2).

Фанетычны мінімізатар **BE EN RU ?**

База:

Адзінка:

Мяжа пошуку:

[СПАМПАВАЦЬ](#) мінімізаваны корпус сказаў — 847.

[СПАМПАВАЦЬ](#) спіс унікальных фанетычных адзінак — 667.

Мал. 2. Карыстальніцкі інтэрфейс вэб-сэрвіса «Фанетычны мінімізатар»

Падрыхтоўка тэкставага корпуса для правядзення эксперыменту. Для правядзення эксперыменту па мінімізацыі вялікага аб'ёму беларускіх тэкстаў быў узяты дампы беларускага раздзела Вікіпедыі за 1 красавіка 2018 года (URL: <https://dumps.wikimedia.org/bewiki/20180401/bewiki-20180401-pages-articles.xml.bz2>). Агульны аб'ём тэкставых дадзеных склаў 191 184 162 сімвалаў.

Дадзенае мноства тэкстаў было разбіта на сказы пры дапамозе рэгулярнага выразу

$$\{ \{ BelUpper \} \} [\{ Bel \} \{ InWords \} \{ Delimiters \}] + [\. \! \? \dots] + ,$$

дзе $\{ BelUpper \}$ – мноства ўсіх вялікіх літар беларускага алфавіту; $\{ Bel \}$ – мноства ўсіх літар беларускага алфавіту; $\{ InWords \}$ – мноства сімвалаў, якія могуць сустракацца ў сярэдзіне слова; $\{ Delimiters \}$ – мноства сімвалаў, якія могуць сустракацца ў сказе паміж словамі, у тым ліку прабельныя сімвалы, знакі прыпынку і інш. Такім чынам дадзены рэгулярны выраз абіраў у якасці сказаў радкі, што пачыналіся з вялікай літары беларускага алфавіту і заканчваліся адным з наступных знакаў прыпынку: кропкай, клічнікам, пыталнікам або шматкроп'ем. Цэла сказа маглі складаць толькі літары беларускага алфавіту, знакі прыпынку, прабельныя сімвалы і сімвалы, якія могуць уваходзіць у склад беларускіх слоў (апострафы, націскі, злучкі). Яшчэ адной умовай абрання сказа з'яўляўся яго аб'ём: сказы ў зыходны корпус прымаліся пры аб'ёме не менш за 5 і не больш за 20 слоў. У выніку з мноства тэкстаў беларускага раздзела Вікіпедыі было вылучана 295 312 сказаў.

Для большай аптымальнасці правядзення эксперыменту атрыманы корпус сказаў быў адразу трансфармаваны ў алафонны запіс пры дапамозе вэб-сінтэзатара маўлення па тэксце (URL: <https://corpus.by/TextToSpeechSynthesizer>), які дае больш за 99 % дакладнасці пры фанетызацыі беларускіх слоў [3]. Мінімальнай фанетычнай адзінкай атрыманага запісу з'яўляецца алафон.

Эксперыментальныя дадзеныя. Пры дапамозе распрацаванага вэб-сэрвіса «Фанетычны мінімізатар» быў праведзены шэраг эксперыменту. Мэтай эксперыменту з'яўлялася вызначэнне параметраў, пры якіх будзе атрымана мінімізаванае фанетычна поўнае мноства сказаў аптымальнага аб'ёму. У якасці зыходнай базы быў выкарыстаны корпус сказаў, атрыманых з беларускага раздзела Вікіпедыі.

Змяненню падвяргаліся наступныя параметры:

- базавая фанетычная адзінка (напрыклад, алафон, дыфон, трыфон і інш.);
- мяжа пошуку – колькасны параметр, да дасягнення якога вядзецца пошук кожнай унікальнай фанетычнай адзінкі (напрыклад, сказы з алафонам N'004 адбіраюцца ў выніковае мноства, пакуль не будзе знойдзена мінімум тры такія алафоны).

У табліцы прыведзены вынікі эксперыменту.

Колькасняя характарыстыкі, атрыманыя пры мінімізацыі тэкставага корпуса

Базавая фанетычная адзінка	Усяго фанетычных адзінак	Колькасць сказаў		
		пры мяжы 1	пры мяжы 2	пры мяжы 3
Алафон (поўны запіс)	667	470	662	847
Дыфон (поўны запіс)	11 846	7498	10 420	13 080
Трыфон (поўны запіс)	79 015	40 502	52 991	63 873
Склад (поўны запіс)	36 823	31 984	41 390	49 463
Алафон (скарочаны запіс)	101	70	105	138
Дыфон (скарочаны запіс)	3181	2117	3009	3821
Трыфон (скарочаны запіс)	31 628	18 076	24 276	29 777
Склад (скарочаны запіс)	18 538	18 024	23 703	28 742

Паводле атрыманых вынікаў можна зрабіць наступныя высновы:

– для стварэння фанетычна поўнага корпуса аб’ёмам у межах некалькіх тысяч сказаў у якасці базавай фанетычнай адзінкі найлепш падыходзіць дыфон (пры скарочаным запісе алафонаў);

– пры патрэбе выкарыстаць буйнейшую базавую адзінку больш мэтазгодна абраць склад, чым трыфон;

– для большай мінімізацыі можна звярнуцца да стварэння фанетычна няпоўнага корпуса, выключаючы са спісу неабходных фанетычных адзінак найбольш рэдкія, у прыватнасці тыя, што сустраляся ў корпусе толькі адзін раз.

Заклучэнне. Зыходнай кропкай для даследавання, прыведзенага ў артыкуле, з’яўляецца ідэя выкарыстання гібрыднай структуры сінтэзатара маўлення ва ўмовах малой забяспечанасці той ці іншай мовы якаснымі і добра структураванымі лінгваакустычнымі рэсурсамі. Пры неабходнасці змяншэння аб’ёму зыходных тэкставых і гукавых дадзеных найбольш сучасныя дасягненні, звязаныя са стварэннем моўных мадэляў і навучаннем нейронных сетак, могуць быць аптымальна аб’яднаны з больш састарэлымі метадамі, уласцівымі канкатэ-натыўнаму падыходу. У кантэксце беларускай мовы прапанаваная структура ўяўляе сабой сумяшчэнне блокаў тэкставай і фанетычнай апрацоўкі, якія выкарыстоўваюць спісы правіл, з блокамі акустычнай і інтанацыйнай апрацоўкі, праца якіх заснавана на корпусным метадзе сінтэзу маўлення. Аднак выкарыстанне такога падыходу зніжае аб’ёмы неабходных зыходных дадзеных не ў дастатковай ступені, што ставіць распрацоўшчыкаў перад патрэбай у дадатковым адборы рэсурсаў.

Распрацаваны аўтарам алгарытм фанетычнай мінімізацыі корпуса тэкстаў на беларускай мове пакліканы аўтаматызаваць працэс далейшага зніжэння аб’ёму дадзеных для навучання сістэмы сінтэзу маўлення. Паводле атрыманых падчас даследавання вынікаў можна зрабіць высновы аб мэтазгоднасці выкарыстання ў якасці базавай фанетычнай адзінкі пры мінімізацыі скарочанага запісу дыфона, адзначыць вартасць выкарыстання складу ў якасці базавай фанетычнай адзінкі, падкрэсліць магчымасць далейшай мінімізацыі пры дапамозе ручной працы, яшчэ большага скарачэння параметраў алафонаў або дапушчэння фанетычнай непаўнаты корпуса праз выключэнне са спісу неабходных фанетычных адзінак найбольш рэдкіх. Разам з тым варта заўважыць, што мінімізацыя аб’ёму зыходных дадзеных не павінна ісці на шкоду не толькі фанетычнай і інтанацыйнай паўнаце корпуса, але і зручнасці працы з корпусам для дыктараў.

Спіс выкарыстаных крыніц

1. Safarik, R. Unified approach to development of ASR systems for east slavic languages / R. Safarik, J. Nouza; ed. N. Camelin, Y. Estève, C. Martín-Vide // Proc. of 5th Intern. Conf. «Statistical Language and Speech Processing» (SLSP’2017), Le Mans, France, 23–25 Oct. 2017. – Springer, 2017. – P. 193–203.

2. Гецэвіч, Ю. С. Аўтаматызаваная апрацоўка сімвальных выразаў у тэкстах для сістэмы сінтэзу беларускага маўлення / Ю. С. Гецэвіч // Информатика. – 2011. – № 4(32). – С. 82–93.

3. Лысы, С. І. Генерацыя нацыянальнай транскрыпцыі тэкстаў на беларускай мове / С. І. Лысы, Ю. С. Гецэвіч // *Информатика*. – 2017. – № 2(54). – С. 84–92.
4. Hunt, A. Unit selection in a concatenative speech synthesis system using a large speech database / A. Hunt, A. Black // *Proc. of IEEE Intern. Conf. «Acoustic, Speech and Signal Processing» (ICASSP'96)*, Atlanta, USA, 7–10 May 1996. – Atlanta, 1996. – Vol. 1. – P. 373–376.
5. Лобанов, Б. М. Компьютерный синтез и клонирование речи / Б. М. Лобанов, Л. И. Цирульник // Минск : Беларус. навука, 2008. – 344 с.
6. Segment selection in the L&H Realspeak laboratory TTS system / G. Coorman [et. al.] // *Proc. of 6th Intern. Conf. «Spoken Language Processing» (ICSLP'2000)*, Beijing, China, 16–20 Oct. 2000. – Beijing, 2000. – Vol. 2. – P. 395–398.
7. Godfrey, J. Language Resources / J. Godfrey, A. Zampolli // *Survey of the State of the Art in Human Language Technology*. – Cambridge University Press, 1996. – Ch. 12. – P. 357–384.
8. Zinovieva, N. Phonetically sufficient allophonic database for concatenation synthesis of russian speech / N. Zinovieva // *Proc. of the 13th Section «Intern. Congress of Phonetic Sciences» (ICPhS'95)*, Stockholm, Sweden, 13–19 Aug. 1995. – Stockholm, 1995. – Vol. 2. – P. 358–361.
9. Fotinea, S.-E. Constructing a segment database for greek time domain speech synthesis / S.-E. Fotinea, G. Tambouratzis, G. Carayannis // *Proc. of 7th European Conf. «Speech Communication and Technology» (EUROSPEECH'2001)*, Aalborg, Denmark, 3–7 Sept. 2001. – Aalborg, 2001. – Vol. 3. – P. 2075–2078.
10. Lambert, T. A database design for a TTS synthesis system using lexical diphones / T. Lambert, A. Breen // *Proc. of 9th European Conf. «Speech Communication and Technology» (InterSpeech'2004)*, Jeju Island, Korea, 4–8 Oct. 2004. – Jeju Island, 2004. – P. 1381–1384.
11. Lyudovyk, T. Speech databases used for concatenative speech synthesis / T. Lyudovyk, M. Sazhok // *Proc. of 7th All-Ukrainian Intern. Conf. on Signal/Image Processing and Pattern Recognition (UkrObraz'2004)*. – Kyjiv, 2004. – P. 111–114.
12. Закревский, А. Д. Основы логического проектирования / А. Д. Закревский, Ю. В. Поттосин, Л. Д. Черемисинова. – Минск : ОИПИ НАН Беларуси, 2004. – Кн. 1. Комбинаторные алгоритмы дискретной математики. – 226 с.
13. *Introduction to Algorithms* / T. H. Cormen [et. al.]. – 3rd ed. – Cambridge : The MIT Press, 2009. – 1292 p.
14. Hue, X. Genetic algorithms for optimization / X. Hue. – Edinburgh : Edinburgh Parallel Computing Centre Press, 1997. – 276 p.
15. Matoušek, J. ARTIC: A new Czech text-to-speech system using statistical approach to speech segment database construction / J. Matoušek, J. Psutka // *Proc. of the 6th Intern. Conf. on Spoken Language Processing (ICSLP'2000)*, Beijing, China, 16–20 Oct. 2000. – Beijing, 2000. – Vol. 4. – P. 612–615.
16. Barbot, N. Comparing performance of different set-covering strategies for linguistic content optimization in speech corpora / N. Barbot, O. Boeffard, A. Delhay // *Proc. of the Intern. Conf. on Language Resources and Evaluation (LREC'12)*. – Istanbul, 2012. – P. 969–974.
17. Development of syllable-based text to speech synthesis system in Bengali / N. P. Narendra [et. al.] // *Intern. J. of Speech Technology*. – 2011. – No. 14(3). – P. 167–181.
18. Kayte, S. A review of unit selection speech synthesis / S. Kayte, M. Mundada, C. Kayte // *Intern. J. of Advanced Research in Computer Science and Software Engineering*. – 2015. – No. 5(10). – P. 475–479.
19. Corpus and voices for catalan speech synthesis / A. Bonafonte [et. al.] // *Proc. of the Intern. Conf. on Language Resources and Evaluation (LREC'2008)*, Marrakech, Morocco, 26 May–1 June 2008. – Marrakech, 2008. – P. 3325–3329.
20. Casademont, E. G. Building synthetic voices in the META-NET framework / E. G. Casademont, A. Bonafonte, M. Moreno // *Proc. of the 8th Intern. Conf. on Language Resources and Evaluation (LREC'12)*, Istanbul, Turkey, 21–27 May 2012. – Istanbul, 2012. – P. 3322–3326.

References

1. Safarik R., Nouza J., ed. Camelin N., Estève Y., Martín-Vide C. Unified approach to development of ASR systems for east slavic languages. *Proceedings of 5th International Conference "Statistical Language and Speech Processing" (SLSP'2017)*, Le Mans, France, 23–25 October 2017. Springer, 2017, pp. 193–203.
2. Hetsevich Yu. S. Аўтаматызаваная апрацоўка сімвалных выразаў у тэкстах для сістэмы сінтэзу біеларускага маўліення [Automated processing symbol expressions in the texts for belarusian speech-to-text synthesis]. *Informatika [Informatics]*, 2011, no. 4(32), pp. 82–93 (in Belarusian).

3. Lysy S. I., Hetsevich Yu. S. Hienieracyja nacyjanal'naj transkrypcyi tekstaŭ na bielaruskaj movie [Generating the national transcription of texts in Belarusian]. *Informatika [Informatics]*, 2017, no. 2(54), pp. 84–92 (in Belarusian).
4. Hunt A., Black A. Unit selection in a concatenative speech synthesis system using a large speech database. *Proceedings of IEEE International Conference "Acoustic, Speech and Signal Processing" (ICASSP '96), Atlanta, USA, 7–10 May 1996*. Atlanta, 1996, vol. 1, pp. 373–376.
5. Lobanov B. M., Cirul'nik L. I. Komp'yuternyj sintez i klonirovanie rechi. *Computer Synthesis and Speech Cloning*. Minsk, Belaruskaya navuka, 2008, 344 p. (in Russian).
6. Coorman G., Fackrell J., Rutten P., Van Coile B. Segment selection in the L&H Realspeak laboratory TTS system. *Proceedings of 6th International Conference "Spoken Language Processing" (ICSLP'2000), Beijing, China, 16–20 October 2000*. Beijing, 2000, vol. 2, pp. 395–398.
7. Godfrey J., Zampolli A. Language Resources. *Survey of the State of the Art in Human Language Technology*. Cambridge University Press, 1996, ch. 12, pp. 357–384.
8. Zinovieva N. Phonetically sufficient allophonic database for concatenation synthesis of russian speech. *Proceedings of the 13th Section "International Congress of Phonetic Sciences" (ICPhS'95), Stockholm, Sweden, 13–19 August 1995*. Stockholm, 1995, vol. 2, pp. 358–361.
9. Fotinea S.-E., Tambouratzis G., Carayannis G. Constructing a segment database for greek time domain speech synthesis. *Proceedings of 7th European Conference "Speech Communication and Technology" (EUROSPEECH'2001), Aalborg, Denmark, 3–7 September 2001*. Aalborg, 2001, vol. 3, pp. 2075–2078.
10. Lambert T., Breen A. A database design for a TTS synthesis system using lexical diphones. *Proceedings of 9th European Conference "Speech Communication and Technology" (InterSpeech'2004), Jeju Island, Korea, 4–8 October 2004*. Jeju Island, 2004, pp. 1381–1384.
11. Lyudovyk T., Sazhok M. Speech databases used for concatenative speech synthesis. *Proceedings of 7th All-Ukrainian International Conference on Signal/Image Processing and Pattern Recognition (UkrObraz'2004)*. Kyjiv, 2004, pp. 111–114.
12. Zakrevskij A. D., Pottosin Yu. V., Cheremisinova L. D. Osnovy logicheskogo proektirovaniya [Basics of logical design]. Kniga 1. Kombinatornye algoritmy diskretnoj matematiki [Book 1. Combinatorial algorithms of discrete mathematics]. Minsk, the United Institute of Informatics Problems National Academy of Sciences of Belarus, 2004, 226 p. (in Russian).
13. Cormen T. H., Leiserson Ch. E., Rivest R. L., Stein C. *Introduction to Algorithms. 3d Ed.* Cambridge, The MIT Press, 2009, 1292 p.
14. Hue X. *Genetic Algorithms for Optimization*. Edinburgh, Edinburgh Parallel Computing Centre Press, 1997, 276 p.
15. Matoušek J., Psutka J. ARTIC: A new Czech text-to-speech system using statistical approach to speech segment database construction. *Proceedings of the 6th International Conference on Spoken Language Processing (ICSLP'2000), Beijing, China, 16–20 October 2000*. Beijing, 2000, vol. 4, pp. 612–615.
16. Barbot N., Boëffard O., Delhay A. Comparing performance of different set-covering strategies for linguistic content optimization in speech corpora. *Proceedings of the International Conference on Language Resources and Evaluation (LREC'12)*. Istanbul, 2012, pp. 969–974.
17. Narendra N. P., Rao K. S., Ghosh K., Vempada R. R., Maity S. Development of syllable-based text to speech synthesis system in Bengali. *International Journal of Speech Technology*, 2011, no 14(3), pp. 167–181.
18. Kayte S., Mundada M., Kayte C. A review of unit selection speech synthesis. *International Journal of Advanced Research in Computer Science and Software Engineering*, 2015, no 5(10), pp. 475–479.
19. Bonafonte A., Adell J., Esquerra I., Gallego S., Moreno A., Pérez J. Corpus and voices for catalan speech synthesis. *Proceedings of the International Conference on Language Resources and Evaluation (LREC'2008), Marrakech, Morocco, 26 May–1 June 2008*. Marrakech, 2008, pp. 3325–3329.
20. Casademont E. G., Bonafonte A., Moreno M. Building synthetic voices in the META-NET framework. *Proceedings of the 8th International Conference on Language Resources and Evaluation (LREC'12), Istanbul, Turkey, 21–27 May 2012*. Istanbul, 2012, pp. 3322–3326.

Інфармацыя пра аўтара

Лысы Станіслаў Іосіфавіч, малодшы навуковы супрацоўнік, Аб'яднаны інстытут праблем інфарматыкі Нацыянальнай акадэміі навук Беларусі, Мінск, Беларусь.
E-mail: stanislau.lysy@gmail.com

Information about the author

Stanislau I. Lysy, Junior Researcher, The United Institute of Informatics Problems of the National Academy of Sciences of Belarus, Minsk, Belarus.
E-mail: stanislau.lysy@gmail.com